

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT: Rodier, Patricia M.
Ingram, Jennifer L.
Figlewicz, Denise A.
Hyman, Susan L.
Stodgell, Christopher J.

(ii) TITLE OF INVENTION: GENETIC POLYMORPHISMS WHICH ARE
ASSOCIATED WITH AUTISM SPECTRUM DISORDERS

(iii) NUMBER OF SEQUENCES: 10

(iv) CORRESPONDENCE ADDRESS:

(A) ADDRESSEE: Nixon, Hargrave, Devans & Doyle LLP
(B) STREET: Clinton Square, P.O. Box 1051
(C) CITY: Rochester
(D) STATE: New York
(E) COUNTRY: U.S.A.
(F) ZIP: 14603-1051

(v) COMPUTER READABLE FORM:

(A) MEDIUM TYPE: Floppy disk
(B) COMPUTER: IBM PC compatible
(C) OPERATING SYSTEM: PC-DOS/MS-DOS
(D) SOFTWARE: PatentIn Release #1.0, Version #1.30

(vi) CURRENT APPLICATION DATA:

(A) APPLICATION NUMBER:
(B) FILING DATE:
(C) CLASSIFICATION:

(vii) PRIOR APPLICATION DATA:

(A) APPLICATION NUMBER: US 60/049,803
(B) FILING DATE: 17-JUN-1997

(viii) ATTORNEY/AGENT INFORMATION:

(A) NAME: Goldman, Michael L.
(B) REGISTRATION NUMBER: 30,727
(C) REFERENCE/DOCKET NUMBER: 176/60181

(ix) TELECOMMUNICATION INFORMATION:

(A) TELEPHONE: (716) 263-1304
(B) TELEFAX: (716) 263-1600

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1008 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

ATGGACAATG CAAGAATGAA CTCCTTCCTG GAATACCCCA TACTTAGCAG TGGCGACTCG	60
GGGACCTGCT CAGCCCGAGC CTACCCCTCG GACCATAGGA TTACAACTTT CCAGTCGTGC	120
GCGGTCAGCG CCAACAGTTG CGGCGGCGAC GACCGCTTCC TAGTGGGCAG GGGGGTGCAG	180
ATCGGTTTCGC CCCACCACCA CCACCACCAC CACCATCACC ACCCCAGCC GGCTACCTAC	240
CAGACTTCCG GGAACCTGGG GGTGTCTTAC TCCCACTCAA GTTGTGGTCC AAGCTATGGC	300
TCACAGAACT TCAGTGCGCC TTACAGCCCC TACGCGTTAA ATCAGGAAGC AGACGTAAGT	360
GGTGGGTACC CCCAGTGC GC TCCCGCTGTT TACTCTGGAA ATCTCTCATC TCCCATGGTC	420
CAGCATCACC ACCACCACCA GGGTTATGCT GGGGGCGCGG TGGGCTCGCC TCAATACATT	480
CACCACTCAT ATGGACAGGA GCACCAGAGC CTGGCCCTGG CTACGTATAA TAACTCCTTG	540
TCCCCTCTCC ACGCCAGCCA CCAAGAAGCC TGTCGCTCCC CCGCATCGGA GACATCTTCT	600
CCAGCGCAGA CTTTTGACTG GATGAAAGTC AAAAGAAACC CTCCCAAAC AGGGAAAGTT	660
GGAGAGTACG GCTACCTGGG TCAACCCAAC GCGGTGCGCA CCAACTTCAC TACCAAGCAG	720
CTCACGGAAC TGGAGAAGGA GTTCCAATT CACAAGTACC TGACGCGCGC CCGCAGGGTG	780
GAGATCGCTG CATCCCTGCA GCTCAACGAG ACCCAAGTGA AGATCTGGTT CCAGAACCGC	840
CGAATGAAGC AAAAGAAACG TGAGAAGGAG GGTCTCTTGC CCATCTCTCC GGCCACCCCG	900
CCAGGAAACG ACGAGAAGGC CGAGGAATCC TCAGAGAAGT CCAGCTCTTC GCCCTGCGTT	960
CCTTCCCCGG GGTCTTCTAC CTCAGACACT CTGACTACCT CCCACTGA	1008

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 335 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

Met	Asp	Asn	Ala	Arg	Met	Asn	Ser	Phe	Leu	Glu	Tyr	Pro	Ile	Leu	Ser
1				5					10					15	
Ser	Gly	Asp	Ser	Gly	Thr	Cys	Ser	Ala	Arg	Ala	Tyr	Pro	Ser	Asp	His
			20					25					30		
Arg	Ile	Thr	Thr	Phe	Gln	Ser	Cys	Ala	Val	Ser	Ala	Asn	Ser	Cys	Gly
		35					40					45			

Gly	Asp	Asp	Arg	Phe	Leu	Val	Gly	Arg	Gly	Val	Gln	Ile	Gly	Ser	Pro
50						55					60				
His	His	His	His	His	His	His	His	His	His	Pro	Gln	Pro	Ala	Thr	Tyr
65					70					75					80
Gln	Thr	Ser	Gly	Asn	Leu	Gly	Val	Ser	Tyr	Ser	His	Ser	Ser	Cys	Gly
				85					90					95	
Pro	Ser	Tyr	Gly	Ser	Gln	Asn	Phe	Ser	Ala	Pro	Tyr	Ser	Pro	Tyr	Ala
			100					105					110		
Leu	Asn	Gln	Glu	Ala	Asp	Val	Ser	Gly	Gly	Tyr	Pro	Gln	Cys	Ala	Pro
		115					120					125			
Ala	Val	Tyr	Ser	Gly	Asn	Leu	Ser	Ser	Pro	Met	Val	Gln	His	His	His
	130					135					140				
His	His	Gln	Gly	Tyr	Ala	Gly	Gly	Ala	Val	Gly	Ser	Pro	Gln	Tyr	Ile
145					150					155					160
His	His	Ser	Tyr	Gly	Gln	Glu	His	Gln	Ser	Leu	Ala	Leu	Ala	Thr	Tyr
				165					170					175	
Asn	Asn	Ser	Leu	Ser	Pro	Leu	His	Ala	Ser	His	Gln	Glu	Ala	Cys	Arg
			180					185					190		
Ser	Pro	Ala	Ser	Glu	Thr	Ser	Ser	Pro	Ala	Gln	Thr	Phe	Asp	Trp	Met
		195					200					205			
Lys	Val	Lys	Arg	Asn	Pro	Pro	Lys	Thr	Gly	Lys	Val	Gly	Glu	Tyr	Gly
	210					215					220				
Tyr	Leu	Gly	Gln	Pro	Asn	Ala	Val	Arg	Thr	Asn	Phe	Thr	Thr	Lys	Gln
225					230					235					240
Leu	Thr	Glu	Leu	Glu	Lys	Glu	Phe	His	Phe	Asn	Lys	Tyr	Leu	Thr	Arg
				245					250					255	
Ala	Arg	Arg	Val	Glu	Ile	Ala	Ala	Ser	Leu	Gln	Leu	Asn	Glu	Thr	Gln
			260					265					270		
Val	Lys	Ile	Trp	Phe	Gln	Asn	Arg	Arg	Met	Lys	Gln	Lys	Lys	Arg	Glu
		275					280					285			
Lys	Glu	Gly	Leu	Leu	Pro	Ile	Ser	Pro	Ala	Thr	Pro	Pro	Gly	Asn	Asp
	290					295					300				
Glu	Lys	Ala	Glu	Glu	Ser	Ser	Glu	Lys	Ser	Ser	Ser	Ser	Pro	Cys	Val
305					310					315					320
Pro	Ser	Pro	Gly	Ser	Ser	Thr	Ser	Asp	Thr	Leu	Thr	Thr	Ser	His	
				325					330					335	

(2) INFORMATION FOR SEQ ID NO:3:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 1008 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

ATGGACAATG CAAGAATGAA CTCCTTCCTG GAATACCCCA TACTTAGCAG TGGCGACTCG	60
GGGACCTGCT CAGCCCAGAG CTACCCCTCG GACCATAGGA TTACAACTTT CCAGTCGTGC	120
GCGGTCAGCG CCAACAGTTG CGGCGGCGAC GACCGCTTCC TAGTGGGCAG GGGGGTGCAG	180
ATCGGTTTCGC CCCACCACCA CCACCACCAC CACCATCGCC ACCCCCAGCC GGCTACCTAC	240
CAGACTTCCG GGAACCTGGG GGTGTCCTAC TCCCACTCAA GTTGTGGTCC AAGCTATGGC	300
TCACAGAACT TCAGTGCGCC TTACAGCCCC TACGCGTTAA ATCAGGAAGC AGACGTAAGT	360
GGTGGGTACC CCCAGTGC GC TCCCGCTGTT TACTCTGGAA ATCTCTCATC TCCCATGGTC	420
CAGCATCACC ACCACCACCA GGGTTATGCT GGGGGCGCGG TGGGCTCGCC TCAATACATT	480
CACCACTCAT ATGGACAGGA GCACCAGAGC CTGGCCCTGG CTACGTATAA TAACTCCTTG	540
TCCCCTCTCC ACGCCAGCCA CCAAGAAGCC TGTCGCTCCC CCGCATCGGA GACATCTTCT	600
CCAGCGCAGA CTTTTGACTG GATGAAAGTC AAAAGAAACC CTCCCAAAAC AGGGAAAGTT	660
GGAGAGTACG GCTACCTGGG TCAACCCAAC GCGGTGCGCA CCAACTTCAC TACCAAGCAG	720
CTCACGGAAC TGGAGAAGGA GTTCCAATT CACAAGTACC TGACGCGCGC CCGCAGGGTG	780
GAGATCGCTG CATCCCTGCA GCTCAACGAG ACCCAAGTGA AGATCTGGTT CCAGAACCGC	840
CGAATGAAGC AAAAGAAACG TGAGAAGGAG GGTCTCTTGC CCATCTCTCC GGCCACCCCG	900
CCAGGAAACG ACGAGAAGGC CGAGGAATCC TCAGAGAAGT CCAGCTCTTC GCCCTGCGTT	960
CCTTCCCCGG GGTCTTCTAC CTCAGACACT CTGACTACCT CCCACTGA	1008

(2) INFORMATION FOR SEQ ID NO:4:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 335 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS:
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

Met	Asp	Asn	Ala	Arg	Met	Asn	Ser	Phe	Leu	Glu	Tyr	Pro	Ile	Leu	Ser	1	5	10	15
Ser	Gly	Asp	Ser	Gly	Thr	Cys	Ser	Ala	Arg	Ala	Tyr	Pro	Ser	Asp	His	20	25	30	
Arg	Ile	Thr	Thr	Phe	Gln	Ser	Cys	Ala	Val	Ser	Ala	Asn	Ser	Cys	Gly	35	40	45	
Gly	Asp	Asp	Arg	Phe	Leu	Val	Gly	Arg	Gly	Val	Gln	Ile	Gly	Ser	Pro	50	55	60	
His	His	His	His	His	His	His	His	Arg	His	Pro	Gln	Pro	Ala	Thr	Tyr	65	70	75	80
Gln	Thr	Ser	Gly	Asn	Leu	Gly	Val	Ser	Tyr	Ser	His	Ser	Ser	Cys	Gly	85	90	95	
Pro	Ser	Tyr	Gly	Ser	Gln	Asn	Phe	Ser	Ala	Pro	Tyr	Ser	Pro	Tyr	Ala	100	105	110	
Leu	Asn	Gln	Glu	Ala	Asp	Val	Ser	Gly	Gly	Tyr	Pro	Gln	Cys	Ala	Pro	115	120	125	
Ala	Val	Tyr	Ser	Gly	Asn	Leu	Ser	Ser	Pro	Met	Val	Gln	His	His	His	130	135	140	
His	His	Gln	Gly	Tyr	Ala	Gly	Gly	Ala	Val	Gly	Ser	Pro	Gln	Tyr	Ile	145	150	155	160
His	His	Ser	Tyr	Gly	Gln	Glu	His	Gln	Ser	Leu	Ala	Leu	Ala	Thr	Tyr	165	170	175	
Asn	Asn	Ser	Leu	Ser	Pro	Leu	His	Ala	Ser	His	Gln	Glu	Ala	Cys	Arg	180	185	190	
Ser	Pro	Ala	Ser	Glu	Thr	Ser	Ser	Pro	Ala	Gln	Thr	Phe	Asp	Trp	Met	195	200	205	
Lys	Val	Lys	Arg	Asn	Pro	Pro	Lys	Thr	Gly	Lys	Val	Gly	Glu	Tyr	Gly	210	215	220	
Tyr	Leu	Gly	Gln	Pro	Asn	Ala	Val	Arg	Thr	Asn	Phe	Thr	Thr	Lys	Gln	225	230	235	240
Leu	Thr	Glu	Leu	Glu	Lys	Glu	Phe	His	Phe	Asn	Lys	Tyr	Leu	Thr	Arg	245	250	255	
Ala	Arg	Arg	Val	Glu	Ile	Ala	Ala	Ser	Leu	Gln	Leu	Asn	Glu	Thr	Gln	260	265	270	
Val	Lys	Ile	Trp	Phe	Gln	Asn	Arg	Arg	Met	Lys	Gln	Lys	Lys	Arg	Glu	275	280	285	
Lys	Glu	Gly	Leu	Leu	Pro	Ile	Ser	Pro	Ala	Thr	Pro	Pro	Gly	Asn	Asp	290	295	300	

Glu Lys Ala Glu Glu Ser Ser Glu Lys Ser Ser Ser Ser Pro Cys Val
305 310 315 320

Pro Ser Pro Gly Ser Ser Thr Ser Asp Thr Leu Thr Thr Ser His
325 330 335

(2) INFORMATION FOR SEQ ID NO:5:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1021 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

TGACGCATGG ACTATAATAG GATGAACTCC TTCTTAGAGT ACCCACTCTG TAACCGGGGA 60
CCCAGCGCCT ACAGCGCCCA CAGCGCCCA ACCTCCTTTC CCCCAAGCTC GGCTCAGGCG 120
GTTGACAGCT ATGCAAGCGA GGGCCGCTAC GGTGGGGGGC TGTCCAGCCC TCGGTTTCAG 180
CAGAACTCCG GCTATCCCGC CCAGCAGCCG CCTTCGACCC TGGGGGTGCC CTTCCCCAGC 240
TCCGCGCCCT CGGGGTATGC TCCTGCCGCC TGCAGCCCCA GCTACGGGCC TTCTCAGTAC 300
TACCCTCTGG GTCAATCAGA AGGAGACGGA GGCTATTTTC ATCCCTCGAG CTACGGGGCC 360
CAGCTAGGGG GCTTGTCGGA TGGCTACGGA GCAGGTGGAG CCGGTCCGGG GCCATATCCT 420
CCGCAGCATC CCCCTTATGG GAACGAGCAG ACCGCGAGCT TTGCACCGGC CTATGCTGAT 480
CTCCTCTCCG AGGACAAGGA AACACCCTGC CCTTCAGAAC CTAACACCCC CACGGCCCCG 540
ACCTTCGACT GGATGAAGGT TAAGAGAAAC CCACCAAGA CAGCGAAGGT GTCAGAGCCA 600
GGCCTGGGCT CGCCAGTGG CCTCCGCACC AACTTCACCA CAAGGCAGCT GACAGAACTG 660
GAAAAGGAGT TCCATTTCAA CAAGTACCTG AGCCGGGCCC GGAGGGTGGA GATTGCCGCC 720
ACCCTGGAGC TCAATGAAAC ACAGGTCAAG ATTTGGTTCC AGAACCGACG AATGAAGCAG 780
AAGAAGCGCG AGCGAGAGGG AGGTCGGGTC CCCCAGCCC CACCAGGCTG CCCCAAGGAG 840
GCAGCTGGAG ATGCCTCAGA CCAGTCGACA TGCACCTCCC CGGAAGCCTC ACCCAGCTCT 900
GTCACCTCCT GAACTGAACC TAGCCACCAA TGGGGCTTCC AGGCACTGGA GCGCCCCAGT 960
CCAGCCCTAT CCCAGGCTCT CCCAACCCAG GCCTGGCTTC ACTGCCTGGG ATCTCTAGGC 1020
T 1021

(2) INFORMATION FOR SEQ ID NO:6:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 301 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS:
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

Met Asp Tyr Asn Arg Met Asn Ser Phe Leu Glu Tyr Pro Leu Cys Asn
1 5 10 15
Arg Gly Pro Ser Ala Tyr Ser Ala His Ser Ala Pro Thr Ser Phe Pro
 20 25 30
Pro Ser Ser Ala Gln Ala Val Asp Ser Tyr Ala Ser Glu Gly Arg Tyr
 35 40 45
Gly Gly Gly Leu Ser Ser Pro Ala Phe Gln Gln Asn Ser Gly Tyr Pro
50 55 60
Ala Gln Gln Pro Pro Ser Thr Leu Gly Val Pro Phe Pro Ser Ser Ala
65 70 75 80
Pro Ser Gly Tyr Ala Pro Ala Ala Cys Ser Pro Ser Tyr Gly Pro Ser
 85 90 95
Gln Tyr Tyr Pro Leu Gly Gln Ser Glu Gly Asp Gly Gly Tyr Phe His
100 105 110
Pro Ser Ser Tyr Gly Ala Gln Leu Gly Gly Leu Ser Asp Gly Tyr Gly
115 120 125
Ala Gly Gly Ala Gly Pro Gly Pro Tyr Pro Pro Gln His Pro Pro Tyr
130 135 140
Gly Asn Glu Gln Thr Ala Ser Phe Ala Pro Ala Tyr Ala Asp Leu Leu
145 150 155 160
Ser Glu Asp Lys Glu Thr Pro Cys Pro Ser Glu Pro Asn Thr Pro Thr
165 170 175
Ala Arg Thr Phe Asp Trp Met Lys Val Lys Arg Asn Pro Pro Lys Thr
180 185 190
Ala Lys Val Ser Glu Pro Gly Leu Gly Ser Pro Ser Gly Leu Arg Thr
195 200 205
Asn Phe Thr Thr Arg Gln Leu Thr Glu Leu Glu Lys Glu Phe His Phe
210 215 220
Asn Lys Tyr Leu Ser Arg Ala Arg Arg Val Glu Ile Ala Ala Thr Leu
225 230 235 240
Glu Leu Asn Glu Thr Gln Val Lys Ile Trp Phe Gln Asn Arg Arg Met
245 250 255

Lys Gln Lys Lys Arg Glu Arg Glu Gly Gly Arg Val Pro Pro Ala Pro
260 265 270
Pro Gly Cys Pro Lys Glu Ala Ala Gly Asp Ala Ser Asp Gln Ser Thr
275 280 285
Cys Thr Ser Pro Glu Ala Ser Pro Ser Ser Val Thr Ser
290 295 300

(2) INFORMATION FOR SEQ ID NO:7:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1030 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

TGACGCATGG ACTATAATAG GATGAACTCC TTCTTAGAGT ACCCACTCTG TAACCGGGGA 60
CCCAGCGCCT ACAGCGCCCA CAGCGCCAC AGCGCCCCAA CCTCCTTTCC CCCAAGCTCG 120
GCTCAGGCGG TTGACAGCTA TGCAAGCGAG GGCCGCTACG GTGGGGGGCT GTCCAGCCCT 180
GCGTTTCAGC AGAACTCCGG CTATCCCGCC CAGCAGCCGC CTTGACCCT GGGGGTGCCC 240
TTCCCCAGCT CCGCGCCCTC GGGGTATGCT CTGCGGCCT GCAGCCCCAG CTACGGGCCT 300
TCTCAGTACT ACCCTCTGGG TCAATCAGAA GGAGACGGAG GCTATTTTCA TCCCTCGAGC 360
TACGGGGCCC AGCTAGGGGG CTTGTCCGAT GGCTACGGAG CAGGTGGAGC CGGTCCGGGG 420
CCATATCCTC CGCAGCATCC CCCTTATGGG AACGAGCAGA CCGCGAGCTT TGCACCGGCC 480
TATGCTGATC TCCTCTCCGA GGACAAGGAA ACACCCTGCC CTTCAGAACC TAACACCCCC 540
ACGGCCCCGA CCTTCGACTG GATGAAGGTT AAGAGAAACC CACCCAAGAC AGCGAAGGTG 600
TCAGAGCCAG GCCTGGGCTC GCCCAGTGGC CTCCGCACCA ACTTCACCAC AAGGCAGCTG 660
ACAGAACTGG AAAAGGAGTT CCATTTCAAC AAGTACCTGA GCCGGGCCCCG GAGGGTGGAG 720
ATTGCCGCCA CCCTGGAGCT CAATGAAACA CAGGTCAAGA TTTGGTTCCA GAACCGACGA 780
ATGAAGCAGA AGAAGCGCGA GCGAGAGGGA GGTGCGGTCC CCCCAGCCCC ACCAGGCTGC 840
CCCAAGGAGG CAGCTGGAGA TGCCTCAGAC CAGTCGACAT GCACCTCCCC GGAAGCCTCA 900
CCCAGCTCTG TCACCTCCTG AACTGAACCT AGCCACCAAT GGGGCTTCCA GGCAGTGGAG 960
CGCCCCAGTC CAGCCCTATC CCAGGCTCTC CCAACCCAGG CCTGGCTTCA CTGCCTGGGA 1020
TCTCTAGGCT 1030

(2) INFORMATION FOR SEQ ID NO:8:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 304 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS:
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

Met	Asp	Tyr	Asn	Arg	Met	Asn	Ser	Phe	Leu	Glu	Tyr	Pro	Leu	Cys	Asn	
1				5					10					15		
Arg	Gly	Pro	Ser	Ala	Tyr	Ser	Ala	His	Ser	Ala	His	Ser	Ala	Pro	Thr	
			20					25					30			
Ser	Phe	Pro	Pro	Ser	Ser	Ala	Gln	Ala	Val	Asp	Ser	Tyr	Ala	Ser	Glu	
		35					40					45				
Gly	Arg	Tyr	Gly	Gly	Gly	Leu	Ser	Ser	Pro	Ala	Phe	Gln	Gln	Asn	Ser	
	50					55					60					
Gly	Tyr	Pro	Ala	Gln	Gln	Pro	Pro	Ser	Thr	Leu	Gly	Val	Pro	Phe	Pro	
65					70					75					80	
Ser	Ser	Ala	Pro	Ser	Gly	Tyr	Ala	Pro	Ala	Ala	Cys	Ser	Pro	Ser	Tyr	
				85					90					95		
Gly	Pro	Ser	Gln	Tyr	Tyr	Pro	Leu	Gly	Gln	Ser	Glu	Gly	Asp	Gly	Gly	
			100					105					110			
Tyr	Phe	His	Pro	Ser	Ser	Tyr	Gly	Ala	Gln	Leu	Gly	Gly	Leu	Ser	Asp	
		115					120					125				
Gly	Tyr	Gly	Ala	Gly	Gly	Ala	Gly	Pro	Gly	Pro	Tyr	Pro	Pro	Gln	His	
	130					135					140					
Pro	Pro	Tyr	Gly	Asn	Glu	Gln	Thr	Ala	Ser	Phe	Ala	Pro	Ala	Tyr	Ala	
145					150					155					160	
Asp	Leu	Leu	Ser	Glu	Asp	Lys	Glu	Thr	Pro	Cys	Pro	Ser	Glu	Pro	Asn	
				165					170					175		
Thr	Pro	Thr	Ala	Arg	Thr	Phe	Asp	Trp	Met	Lys	Val	Lys	Arg	Asn	Pro	
			180					185					190			
Pro	Lys	Thr	Ala	Lys	Val	Ser	Glu	Pro	Gly	Leu	Gly	Ser	Pro	Ser	Gly	
		195					200					205				
Leu	Arg	Thr	Asn	Phe	Thr	Thr	Arg	Gln	Leu	Thr	Glu	Leu	Glu	Lys	Glu	
	210					215					220					
Phe	His	Phe	Asn	Lys	Tyr	Leu	Ser	Arg	Ala	Arg	Arg	Val	Glu	Ile	Ala	
225					230					235					240	

Ala	Thr	Leu	Glu	Leu	Asn	Glu	Thr	Gln	Val	Lys	Ile	Trp	Phe	Gln	Asn	
				245					250					255		
Arg	Arg	Met	Lys	Gln	Lys	Lys	Arg	Glu	Arg	Glu	Gly	Gly	Arg	Val	Pro	
			260					265					270			
Pro	Ala	Pro	Pro	Gly	Cys	Pro	Lys	Glu	Ala	Ala	Gly	Asp	Ala	Ser	Asp	
		275					280					285				
Gln	Ser	Thr	Cys	Thr	Ser	Pro	Glu	Ala	Ser	Pro	Ser	Ser	Val	Thr	Ser	
	290					295					300					

(2) INFORMATION FOR SEQ ID NO:9:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

- (A) DESCRIPTION: /desc = "primer"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

GCATGGACTA TAATAGGATG

20

(2) INFORMATION FOR SEQ ID NO:10:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

- (A) DESCRIPTION: /desc = "primer"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

TCTTGGGTGG GTTTCTCTTA

20